*European Society of Computational Methods in Sciences and Engineering (ESCMSE)*

*Journal of Numerical Analysis, Industrial and Applied Mathematics (JNAIAM)*
vol. 4, no. 1-2, 2009, pp. 1-9
ISSN 1790–8140

# Forty-Five Years of A-stability[1]

**J. C. Butcher**[2]

Department of Mathematics,
University of Auckland,
New Zealand

Received 2 March, 2009; accepted in revised form 12 April, 2009

*Abstract:* We discuss two events with profound implications on the way initial value problems are solved numerically. The first was the identification of stiffness as a widely spread phenomenon affecting the ability to obtain useful results. The second was the definition of A-stability as an important approach to overcoming the effect of stiffness. Not only was the idea associated with A-stability significant in its own time but it has had long term effects including new theoretical questions as well as the tools for solving them.



*Keywords:* Stiff problems, A-stability, stability barriers, order stars, order arrows, non-linear stability.

*Mathematics Subject Classification:* 65L05, 65L06

## 1 Introduction

To misquote Tolstoy, "All easy problems are much the same but every difficult problem is difficult in its own way". However, there are some common features to some difficult problems and, in 1952, Curtiss and Hirschfelder [5] identified a certain type of difficulty and gave it the name stiffness. Many examples were recognised in physics and engineering problems and now stiffness is seen everywhere! One early reaction to the existence of stiff problems was the definition by Dahlquist [6] in 1963 of A-stability. Although A-stable methods cannot solve all stiff problems, they can solve many of them; furthermore a method which is not A-stable will encounter some level of difficulty with any stiff problem.

Throughout the history of A-stability, various attempts have been made to determine which method types are or are not A-stable. This has led to a number of broad-based results and a similar number of conjectured propositions some of which eventually became settled. In parallel with these developments attempts were made to weaken the definition of A-stability on the one hand and strengthen it on the other. Weakening the definition leads to the inclusion of some potentially efficient numerical methods as being of some practical value and strengthening the definition will lead to stricter requirements which might be needed to solve some problems.

In Section 2 an informal description of stiffness will be given in terms of a simple example problem. The concept of A-stability, together with a statement of some of the classical order

[1]Published electronically June 15, 2009

[2]Corresponding author. E-mail: butcher@math.auckland.ac.nz

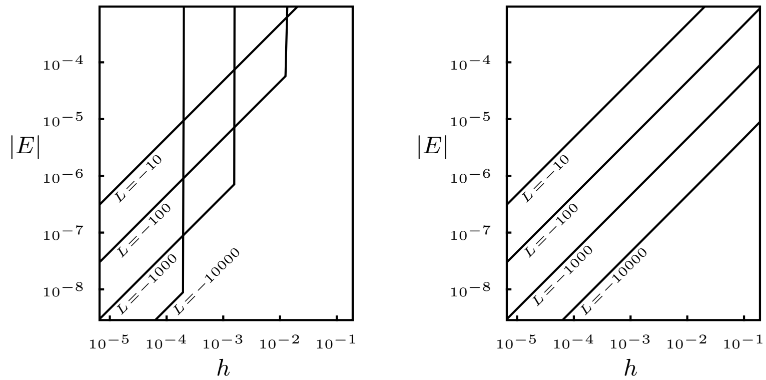

Figure 1: Attempts to solve (1), left: Euler method, right: implicit Euler method

barriers, will be introduced in Section 3. Order stars and order arrows will be surveyed in Section 4, followed in 5 by an outline of their applications to some of the classic order barriers. A brief discussion of weakened stability criteria will be given in Section 6 and this is followed in Section 7 by a brief survey of non-linear stability.

## 2 Stiff problems

The following simple problem is a special case of the differential equation introduced in [12] in a study of the behaviour of Runge–Kutta methods.

$$\frac{dy}{dx} = L\big(y - \sin(x)\big) + \cos(x), \quad y(0) = 0, \tag{1}$$

where $L$ is a very large negative real constant. The exact solution is $y = \sin(x)$.

If we try to solve this problem using the Euler method, the "local truncation error" in step number $n$, when the stepsize is $h$, is approximately $\sin(x_n)h^2/2$. Hence, it should be possible to solve this problem with an error contribution in each step of no more than $\epsilon$, if $h$ is chosen as approximately $\sqrt{2\epsilon}$. The Euler method works very well if $L$ is close to zero, but is virtually unusable if $|L|$ is large. This is shown in the left-hand diagram in Figure 1. Contrasting results for the implicit Euler method are shown in the right-hand diagram. It is seen that satisfactory results are found for a large range of stepsizes for all chosen values of $L$.

Although it is not convenient to attempt a precise characterisation of stiffness, the problem (1) when $-L$ is large, is regarded as stiff because

> The ability to obtain useful answers with the Euler method is limited by stability rather than accuracy.

This gives a good guide to stiffness behaviour in general but, in the case of the example problem, we can understand it better by stripping away the non-linear terms. This leaves us with

$$\frac{dy}{dx} = Ly.$$

The approximation found by the Euler method is

$$y_n = (1 + hL)^n y_0 \approx \exp\big(L(x_n - x_0)\big)y_0,$$

where $nh = x_n - x_0$.

If $hL < -2$, the sequence of $y_n$ values is computed using the *unstable* difference equation

$$y_n = (1 + hL)y_{n-1}. \tag{2}$$

However, if the implicit Euler method is used, we obtain the approximation

$$y_n = (1 - hL)^{-n}y_0 \approx \exp\big(L(x_n - x_0)\big)y_0,$$

found from the *stable* difference equation

$$y_n = (1 - hL)^{-1}y_{n-1}.$$

Note that it would have done no good to use a more accurate explicit method, such as one of the high order Runge-Kutta methods, because $1+hL$ in (2) would then be replaced by a higher degree polynomial in $hL$. Again high values of $hL$ would lead to divergent sequences of $y$ values.

## 3 A-stability

The sequence of approximations over many time-steps satisfies a difference equation and, for linear problems, this is a linear difference equation. If the differential equation is $y' = Ly$ and we write $z = hL$, where $h$ is the stepsize, we ask for what values of $z$ the linear difference equation has only bounded solutions. If this is the case, we say that "$z$ lies in the stability region".

A numerical method is A-stable if all points in the left half of the complex plane belong to the stability region. Suppose a numerical method is applied to a linear problem whose solution is bounded as time increases. The significance of the method being A-stable is that the numerical approximations will also remain bounded as time increases.

We already know the existence of one A-stable method, the implicit Euler method, and one example of a method which is not A-stable, the explicit Euler method. It is known that no genuine explicit method can be A-stable. On the other hand, implicit Runge–Kutta methods can be A-stable and, moreover, can be A-stable with arbitrarily high orders. For linear multistep methods it is an unfortunate fact that A-stable methods cannot have order greater than 2. It is also possible to construct "general linear methods" which are neither Runge–Kutta nor linear multistep. Like Runge–Kutta methods, these are also capable of both high orders and A-stability.

Associated with the search for high order A-stable methods, are a number of mathematical results known as barriers. These express broad ranging conditions under which A-stability is impossible. Some examples are:

1. An A-stable one-stage method, that is a method which involves only a single function evaluation per time step, cannot have order greater than 2. This is known as Dahlquist's second barrier and applies in particular to the BDF methods.

2. An A-stable $s$ stage method cannot have order greater than $2s$. This was formerly known as the Daniel–Moore conjecture [9], until it was proved using order stars [14].

3. An A-stable $s$ stage method whose stability function is a Padé approximation to the exponential function cannot have order less than $2s - 2$. This result, formerly known as the Ehle conjecture [10], was also proved in [14] using order stars.

4. An A-stable multivalue method whose stability function is a generalised Padé approximation cannot have order less than $2s - 2$.

   This statement is known as the Butcher-Chipman conjecture [2] and was proved in [1].

The proofs of these barriers make use of order stars, or the closely related order arrows.

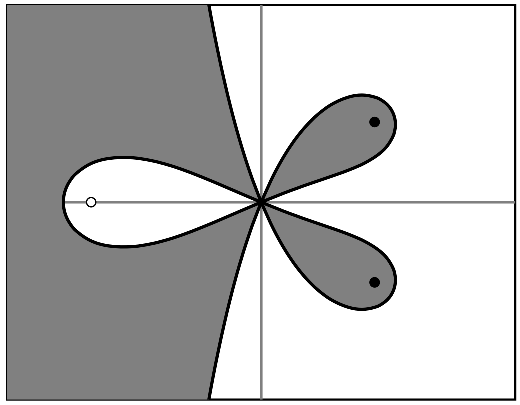

Figure 2: Order star for $R(z) = (1 + \frac{1}{3}z)/(1 - \frac{2}{3}z + \frac{1}{6}z^2)$

## 4 Order stars and order arrows

For a stability function $R(z)$ to correspond to an A-stable numerical method, it is necessary and sufficient that (a) All poles are in the right half-plane and (b) $|R(iy)| \le 1$ for all real $y$.

In [14] Hairer, Nørsett and Wanner proposed considering the regions where $\exp(-z)R(z)$, rather than $R(z)$, lies in the unit disc. Note that the necessary and sufficient conditions (a) and (b) remain intact when this change is made.

But there is a bonus in using $\widetilde{R}(z) = \exp(-z)R(z)$ instead of $R(z)$, because

$$\exp(-z)R(z) = 1 - Cz^{p+1} + O(z^{p+2}).$$

We will illustrate this idea by plotting the "order star" for the function $R(z) = (1 + \frac{1}{3}z)/(1 - \frac{2}{3}z + \frac{1}{6}z^2)$; this is shown in Figure 2. For this function, $p = 3$ and $C = \frac{1}{72}$.

The shaded region represents the set of points for which $|\widetilde{R}(z)| > 1$ and evidently this region does not intersect the imaginary axis. Hence (b) for this function is satisfied and, because (a) is also true, the approximation corresponds to an A-stable method.

To this diagram we can add additional contour lines for $\widetilde{R}(z)$ as well as the orthogonal set of contour lines for $\arg \widetilde{R}(z)$. The new diagram is shown in Figure 3. Now remove all except the lines for which $\arg \widetilde{R}(z) = 0$; that is the lines for which $\widetilde{R}(z)$ is real and positive. This is shown in Figure 4. The arrows emanating from zero are "up-arrows" in which $\widetilde{R}(z) \ge 0$ and "down-arrows" in which $\widetilde{R}(z) \le 0$. Because $\widetilde{R}(z) = 1 - Cz^{p+1} + O(z^{p+2})$, the up and down arrows alternate and their tangents at zero make constant angles $\pi/(p+1)$ between each arrow of one type and the next arrow of the other type. This is illustrated in Figure 5 where the tangents are added to Figure 4. Each of Figures 2, 3 and 4 previously appeared in [1].

The value of $\widetilde{R}(z)$ increases as an up-arrow is traced out from zero and eventually it terminates at a pole or at $-\infty$. With a down-arrow, $\widetilde{R}(z)$ decreases until it terminates at a zero or at $+\infty$.

## 5 Classical barrier results

The second Dahlquist barrier states that an A-stable method cannot have order greater than 2. An order arrow proof of this famous result makes use of the fact that only one up-arrow terminates at a pole because there is only one pole. This means that there must be an up-arrow emanating

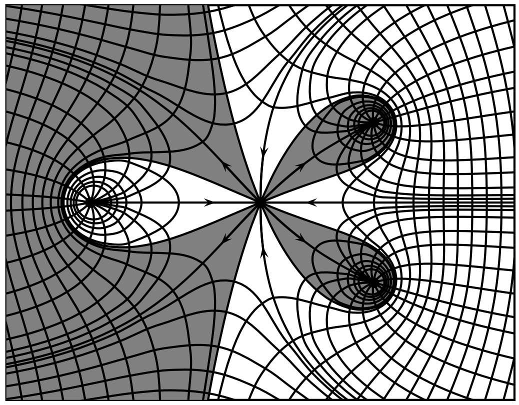

Figure 3: Contour lines for $|\widetilde{R}(z)|$ and $\arg \widetilde{R}(z)$

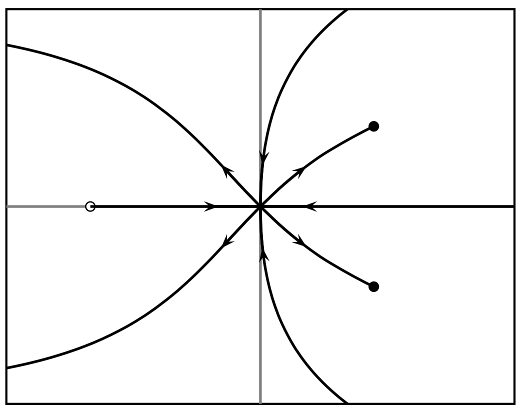

Figure 4: Order arrows for $R(z) = (1 + \frac{1}{3}z)/(1 - \frac{2}{3}z + \frac{1}{6}z^2)$

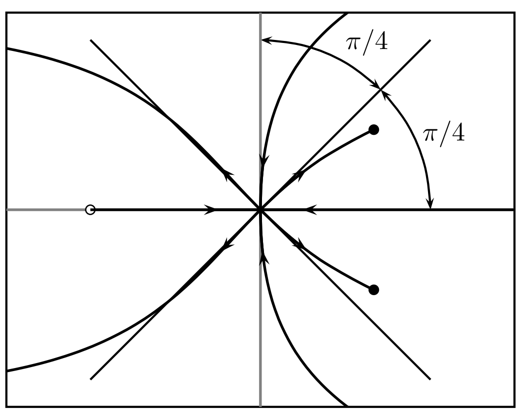


Figure 5: Tangents of arrows at zero.

from zero with argument $2\pi/(p+1)$. Because this terminates at $-\infty$ it must be tangential to, or cross the imaginary axis, unless $p \le 2$.

The result, formerly known as the Daniel–Moore conjecture, states that A-stability is possible for an $s$-stage method only if $p \le 2s$. This was provided with an extremely elegant proof in [14]. However, it can also be proved using order arrows. In this case, for arg $\in [-\pi/2, \pi/2]$, there are at least $[(p+1)/2]$ up-arrows emanating from zero. Of these arrows only $s$ terminate at poles and the remainder must terminate at $-\infty$ and therefore cross or are tangential to the imaginary axis.

Another success of the order star approach was the proof of the result formerly known as the Ehle conjecture. A Runge–Kutta method whose stability function has numerator of degree $n$ and denominator of degree $d$ has maximum possible order $n+d$. If this order is actually achieved, the stability function is a Padé approximation to the exponential function. This stability function satisfies the conditions for A-stability if and only if $n \le d \le n+2$. The difficult part of this result is the statement that A-stability is impossible if $d > n+2$ and is the part put forward as a conjecture by Ehle. The order arrow proof of this result is based on the observation that $d+n+1$ up-arrows alternate with $n+d+1$ down-arrows emanating from zero. Because up-arrows terminate at a pole or at $-\infty$ and the down-arrows terminate at a zero or at $+\infty$, every pole is connected to zero by an up-arrow and every zero is connected to zero by a down-arrow, since otherwise a down-arrow would cross an up-arrow, and this is impossible. For an A-stable method, the $d$ up-arrows which terminate at the poles must all leave zero in the sector $\arg(z) \in (-\pi/2, \pi/2)$. Hence, $2\pi(d-1)/(n+d+1) < 2\pi$, implying that $d \le n+2$.

## 6 Generalizations of A-stability

Because many otherwise satisfactory numerical methods are not A-stable, it is natural to ask if a weakening of this property can still lead to useful methods for the solution of stiff problems. The two most important of these generalized stability definitions are A($\alpha$) [15] and stiff stability [11]. They are both motivated by the same idea, that many stiff problems have spectra which lie in the left half-plane but not near the imaginary axis.

A method is stiffly stable with stiffness abscissa $D$ if the stability region includes all complex numbers $z$ such that $\mathrm{Re}(z) \le -D$. A method is A($\alpha$) stable if the stability region includes all complex numbers $z$ such that $-(\pi-\alpha) \le \arg(z) \le \pi-\alpha$.

We illustrate both these definitions in Figure 6 in the case of the BDF4 method. For this method, $\alpha = 73.352°$ and $D = 0.66667$.

## 7 Non-linear stability

In 1975 Germund Dahlquist introduced the non-linear test problem [7]

$$y'(x) = f(x, y(x)), \quad \langle u - v, f(x,u) - f(x,v)\rangle \le 0. \tag{3}$$

This is a useful model for studying non-linear stiffness because, for this problem,

$$\begin{aligned} \tfrac{d}{dx}\|y(x)-z(x)\|^2 &= 2\,\langle y(x)-z(x), f(x,y(x)-f(x,z(x))\rangle \\ &\le 0, \end{aligned}$$

for two distinct solutions to (3).

For convenience we will replace (3) by a closely related test problem

$$y'(x) = f(x, y(x)), \quad \langle u, f(x,u)\rangle \le 0.$$

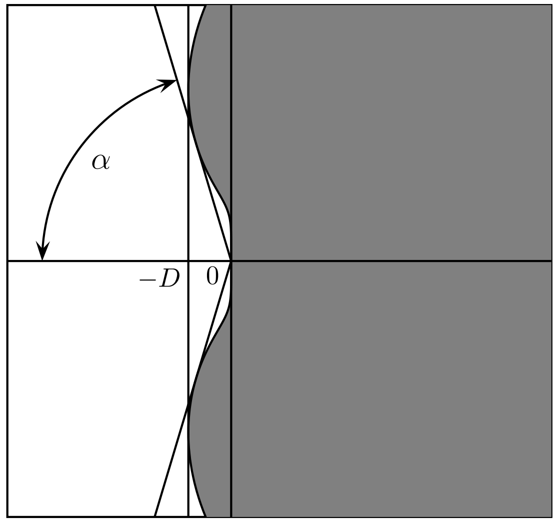


Figure 6: A(α) and stiff stability

Because linear multistep methods are multivalue, it is necessary to express numerical adherence to the non-increasing nature of $\|y(x)\|$ in a non-trivial way. Dahlquist did this by introducing a positive-definite matrix $G$ and defining "G-stability", to mean that

$$\sum_{i,j=1}^{k} g_{ij}\langle y_{n+1-i}, y_{n+1-j}\rangle \leq \sum_{i,j=1}^{k} g_{ij}\langle y_{n-i}, y_{n-j}\rangle.$$

The conditions for this to hold depend on the coefficients $\alpha_0, \alpha_1, \ldots, \alpha_k, \beta_0, \beta_1, \ldots, \beta_k$ in the linear multistep method

$$\alpha_0 y_n + \alpha_1 y_{n-1} + \cdots + \alpha_k y_{n-k} = h\beta_0 f(y_n) + h\beta_1 f(y_{n-1}) + \cdots + h\beta_k f(y_{n-k}), \tag{4}$$

together with the elements of the matrix $G$. In the case of a general linear multistep method, it is found that the analysis cannot be done directly, and it is necessary to use the one-leg counterpart to (4). This one-leg counterpart can be written as

$$\alpha_0 y_n + \alpha_1 y_{n-1} + \cdots + \alpha_k y_{n-k} = hf(\beta_0 y_n + h\beta_1 y_{n-1} + \cdots + h\beta_k y_{n-k}), \tag{5}$$

where the coefficients are normalised so that $\sum_{i=0}^{k} \beta_i = 1$.

Recently Adrian Hill and I have shown how non-linear stability can be established directly in the context of general linear methods [3]. A famous result of Dahlquist states that A-stability is equivalent to G-stability [8].

For Runge-Kutta methods, there is only a single input and output for each step and the criterion for non-linear stability is more natural:

$$\|y_n\| \leq \|y_{n-1}\|.$$

Recall the tableau for an $s$-stage Runge-Kutta method:

$$\begin{array}{c|c} c & A \\ \hline & b^T \end{array} = \begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \cdots & a_{1s} \\ c_2 & a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & \vdots & & \vdots \\ c_s & a_{s1} & a_{s2} & \cdots & a_{ss} \\ \hline & b_1 & b_2 & \cdots & b_s \end{array}$$

In step number $n$,

$$Y_i = y_{n-1} + h\sum_j a_{ij}F_j, \quad F_i = f(Y_i),$$
$$y_n = y_{n-1} + h\sum_i b_i F_i.$$

Introduce the symmetric $s \times s$ matrix $M$ with elements

$$m_{ij} = b_i a_{ij} + b_j a_{ji} - b_i b_j;$$

that is

$$M = \operatorname{diag}(b)A + A^T \operatorname{diag}(b) - bb^T.$$

It is possible to verify the identity

$$\langle y_n, y_n\rangle = \langle y_{n-1}, y_{n-1}\rangle + 2h\sum_i b_i\langle F_i, Y_i\rangle - h^2\sum_{ij} m_{ij}\langle F_i, F_j\rangle \tag{6}$$

and this leads to the definition: A Runge–Kutta method is “B-stable” if $b_i \geq 0$ $(i = 1, 2, \ldots, s)$ and $M$ is positive semi-definite.

If these conditions are satisfied then we obtain the required behaviour for dissipative problems:

$$\|y_n\| \leq \|y_{n-1}\|$$

because of (6). Although A-stability does not imply B-stability, the closely related concepts of AN and BN-stability are equivalent.

The matrix $M$, which plays a central role in B-stability, has since become of crucial importance in the study of canonical Runge–Kutta methods [4] which respect quadratic invariants. Suppose $\langle y, Qy\rangle$ is a quadratic invariant, so that

$$\langle f(y), Qy\rangle = 0,$$

then, noting that (6) still holds if the inner product $\langle\cdot,\cdot\rangle$ is replaced by $\langle\cdot, Q\cdot\rangle$, we see that, if $M = 0$,

$$\langle y_n, Qy_n\rangle = \langle y_{n-1}, Qy_{n-1}\rangle,$$

indicating that invariance is preserved by numerical approximations.

A closely related result is that if $M = 0$, the method preserves symplectic behaviour for Hamiltonian problems [13].

## References

[1] Butcher J. C., Order and stability of generalized Padé approximations, *Appl. Numer. Math.* **59** (2009), 558–567.

[2] Butcher J. C. and Chipman F. H., Generalized Padé approximations to the exponential function, *BIT*, **32** (1992), 118–130.

[3] Butcher J. C. and Hill A. T, Linear multistep methods as irreducible general linear methods, *BIT*, **46** (2006), 5–19.

[4] Cooper G. J., Stability of Runge-Kutta methods for trajectory problems, *IMA J. Numer. Anal.*, **7** (1987), 1–13.

[5] Curtiss C. F. and Hirschfelder J. O, Integration of stiff equations, *Proc. Nat. Acad. Sci. U.S.A.*, **38** (1952), 235–243.

[6] Dahlquist G., A special stability problem for linear multistep methods, *BIT*, **3** (1963), 27–43.

[7] Dahlquist G., Error analysis for a class of methods for stiff non-linear initial value problems, *Lecture Notes in Math.*, **506** (1976), 60–72.

[8] Dahlquist G., G-stability is equivalent to A-stability, *BIT*, **18** (1978), 384–401.

[9] Daniel J. W. and Moore R. E., *Computation and Theory in Ordinary Differential Equations*, W. H. Freeman, 1970.

[10] Ehle B. L., A-stable methods and Padé approximations to the exponential *SIAM J. Math. Anal.*, **4** (1973),. 671–680.

[11] Gear C. W., Numerical Initial Value Problems in Ordinary Differential Equations, Prentice-Hall, Englewood Cliffs, New Jersey, 1971.

[12] Prothero A. and Robinson A., On the stability and accuracy of one-step methods for solving stiff systems of ordinary differential equations, *Math. Comp.*, **28** (1974), 145–162.

[13] Sanz-Serna, J. M., Runge–Kutta schemes for Hamiltonian systems, *BIT*, **28** (1988), 877–883.

[14] Wanner G., Hairer E. and Nørsett S. P., Order stars and stability theorems, *BIT*, **18** (1978), 475–489.

[15] Widlund O. B., A note on unconditionally stable linear multistep methods, *BIT* **7** (1967), 65–70.